



sequence-976414[1].3(1).Oct02.ST25
SEQUENCE LISTING

<110> RHEE, Sang-Ki
SONG, Ki-Bang
KIM, Chul-Ho
RYU, Eun-Ja
LEE, Yong-Bok

<120> Enzymatic production of difructose dianhydride IV from sucrose and relevant enzymes and genes coding for them

<130> 24679

<140> US 09/868,328

<141> 2001-06-18

<150> PCT/KR00/01183

<151> 2000-10-19

<160> 4

<170> PatentIn version 3.1

<210> 1

<211> 521

<212> PRT

<213> Arthrobacter ureafaciens K2032

<400> 1

Met Thr Pro Ala Ile Ser Arg Arg Ala Val Leu Gln Gly Ala Gly Ala
1 5 10 15

Gly Ala Leu Ala Leu Ile Phe Gly Gly Ala Val Pro Pro Ala Ala Arg
20 25 30

Ala Ser Ala Pro Gly Ser Leu Arg Ala Val Tyr His Met Thr Pro Pro
35 40 45

Ser Gly Trp Leu Cys Asp Pro Gln Arg Pro Val Thr Thr His Gly Ala
50 55 60

Tyr Gln Leu Tyr Tyr Leu His Ser Asp Gln Asn Asn Gly Pro Gly Gly
65 70 75 80

Trp Asp His Ala Ser Thr Thr Asp Gly Val Ala Phe Thr His His Gly
85 90 95

Thr Val Met Pro Leu Arg Pro Asp Phe Pro Val Trp Ser Gly Ser Ala
100 105 110

Val Val Gly Thr Ala Asn Thr Ala Gly Phe Gly Ala Gly Ala Val Val
115 120 125

Ala Leu Ala Thr Gln Pro Thr Asp Gly Val Arg Lys Tyr Gln Glu Gln
Page 1

130

135

140

Tyr Leu Tyr Trp Ser Thr Asp Gly Gly Phe Thr Phe Thr Ala Leu Pro
145 150 155 160

Asp Pro Val Ile Val Asn Thr Asp Gly Arg Ala Ala Thr Thr Pro Ala
165 170 175

Glu Ile Glu Asn Ala Glu Trp Phe Arg Asp Pro Lys Ile His Trp Asp
180 185 190

Thr Ala Arg Gly Glu Trp Val Cys Val Ile Gly Arg Leu Arg Tyr Ala
195 200 205

Ala Phe Tyr Thr Ser Pro Asn Leu Arg Asp Trp Thr Leu Arg Arg Asn
210 215 220

Phe Asp Tyr Pro Asn His Ala Leu Gly Gly Ile Glu Cys Pro Asp Leu
225 230 235 240

Phe Glu Ile Thr Ala Asp Asp Gly Thr Arg His Trp Val Leu Ala Ala
245 250 255

Ser Met Asp Ala Tyr Gly Ile Gly Leu Pro Met Thr Tyr Ala Tyr Trp
260 265 270

Thr Gly Thr Trp Asp Gly Glu Gln Phe His Ala Asp Asp Leu Thr Pro
275 280 285

Gln Trp Leu Asp Trp Gly Trp Asp Trp Tyr Ala Ala Val Thr Trp Pro
290 295 300

Ser Ile Asp Ala Pro Glu Thr Lys Arg Leu Ala Ile Ala Trp Met Asn
305 310 315 320

Asn Trp Lys Tyr Ala Ala Arg Asp Val Pro Thr Asp Ala Ser Asp Gly
325 330 335

Tyr Asn Gly Gln Asn Ser Ile Val Arg Glu Leu Arg Leu Ala Arg Gln
340 345 350

Pro Gly Gly Trp Tyr Thr Leu Leu Ser Thr Pro Val Ala Ala Leu Thr
355 360 365

Asn Tyr Val Thr Ala Thr Thr Thr Leu Pro Asp Arg Thr Val Asp Gly
370 375 380

sequence-976414[1].3(1).Oct02.ST25

Ser Ala Val Leu Pro Trp Asn Gly Arg Ala Tyr Glu Ile Glu Leu Asp
385 390 395 400

Ile Ala Trp Asp Thr Ala Thr Asn Val Gly Ile Ser Val Gly Arg Ser
405 410 415

Pro Asp Gly Thr Arg His Thr Asn Ile Gly Lys Tyr Gly Ala Asp Leu
420 425 430

Tyr Val Asp Arg Gly Pro Ser Asp Leu Ala Gly Tyr Ser Leu Ala Pro
435 440 445

Tyr Ser Arg Ala Ala Ala Pro Ile Asp Pro Gly Ala Arg Ser Val His
450 455 460

Leu Arg Ile Leu Val Asp Thr Gln Ser Val Glu Val Phe Val Asn Ala
465 470 475 480

Gly His Thr Val Leu Ser Gln Gln Val His Phe Ala Glu Gly Asp Thr
485 490 495

Gly Ile Ser Leu Tyr Thr Asp Gly Gly Pro Ala His Phe Thr Gly Ile
500 505 510

Val Val Arg Glu Ile Gly Gln Ala Ile
515 520

<210> 2
<211> 2001
<212> DNA
<213> Arthrobacter ureafaciens K2032

<400> 2
gcggtgcacc cgcacttccc ctcgacgacc accgtcccc taccggccga cgcgccggcc 60
cgactgctcc tcagcctaga cgggccctc ctcgaggtct tcgtcgggga cggtgaggcg 120
actgctcga acctggtcct cctggggggcc ggcggtgtga ccgcgagcct cgagacggca 180
cggccaggaa ccgtgcacgt gaccgcgatc gacgtcgagg cgcacagcga tgctgacgcc 240
cctgaacctg ccgccgttct gggctgacga gcgctccac cccgacagct ctccttctac 300
cgctgcccga accaggggtg acgcttcgtc gcgcccaccc gtccacgaga ggaaccagca 360
atgacgccgg ccatctcacg ccgcgccgtg ctccaggag ccggcgccgg agcactcgcc 420
ctgatcttcg gcggtgctgt gccgcctgca gccggggcat ccgctccggg ctcgctccgt 480
gccgtctacc acatgacgcc cccagcggc tggtctctcg accccaacg cccggtcacc 540
accacggcg cctaccagct gtactacctg cactccgacc agaacaacgg ccccggcggc 600
tgggaccacg cgagcacgac cgacggcgtc gccttcacgc accacggcac cgtgatgccg 660

sequence-976414[1].3(1).Oct02.ST25

ctgcggccccg acttccccgt gtggtccggg tcggcggtcg tcggcaccgc gaacacggca	720
gggttcggcg ccggcgcggt cgtcgcgctc gcgaccagc cgaccgacgg cgtccgcaag	780
taccaggagc agtacctcta ctggtcgacc gacggcggggt tcacgttcac cgccctgccc	840
gaccccgta tcgtcaacac cgacggtcgc gccgccacca cgcccgcga gatcgagaac	900
gccgagtggg tccgcgaccc caagatccac tgggacaccg cccgcggaga atgggtctgc	960
gtcatcggac gactgcggta cgccgcgttc tacacctgc cgaacctgc cgactggaca	1020
cttcgccgca acttcgacta cccgaaccac gccctcggcg gcatcgagt ccccgacctg	1080
ttcgagatca ccgcagacga cgggacacgc cactgggtgc tcgccgccag catggacgcc	1140
tacggcatcg gcctccccat gacgtacgcc tactggacag gcacctggga cggcgagcag	1200
ttccacgccg acgacctcac cccgcaatgg ctcgactggg gctgggactg gtacgcggcc	1260
gtcacctggc catcgatcga cgcgccgag accaagcgcc tcgccatcgc gtggatgaac	1320
aactggaagt acgccgcacg cgacgtcccc accgacgcat ccgacggcta caacgggcag	1380
aactcgatcg tccgcgagct gcggctcgcc cgacagcctg gcggctggta caccctcctg	1440
agcacccccg tggcagcgct gacgaactac gtcaccgcca ccaccacact ccccgaccgg	1500
accgtcgacg gcagcgccgt cctgccatgg aacggacgcg catacgagat cgagctcgac	1560
atcgctggg acaccgcgac gaacgtcggc atctcgggtg gccgtcccc cgacggaacc	1620
cggcacacga acatcgga gtagcgagca gacctgtacg tcgaccgagg accctccgac	1680
ctcgccgggt actcgctcgc cccctactcg cgagccgccg ccccatcga ccccggcgcc	1740
cgatccgtgc acctgcgat cctcgtcgac acccagagcg tcgaggcttt cgtcaacgcc	1800
ggccacaccg tgctctcca gcaggtccac ttcgccgagg gcgacacggg aatctcgtc	1860
tacaccgacg gcggccccgc acattcacc ggcacgtcg tccgcgagat tggccaggcg	1920
atctaggcga tgcacaccac accgctcacc gaagccgcgc cccgggagac gacggccgac	1980
aatcgacacg tcctcgtcgt t	2001

<210> 3
 <211> 2001
 <212> DNA
 <213> *Arthrobacter ureafaciens* K2032

<220>
 <221> CDS
 <222> (361)..(1917)
 <223>

<400> 3	
gcggtgcacc ccgacttccc ctcgacgacc accgtcccc taccggccga ccgcccggcc	60
cgactgctcc tcagcctaga cgggcccctc ctcgaggtct tcgtcgggga cggtgaggcg	120

sequence-976414[1].3(1).Oct02.ST25

actgcgtcga acctgggtcct cctggggggcc ggcggtgtga cgcgagcct cgagacggca	180
cggccaggaa ccgtgcacgt gaccgcatc gacgtcgagg cgcccagcga tgctgacgcc	240
cctgaacctg ccgccgttct gggctgacga gcgctccac cccgacagct ctccttctac	300
cgctgcccga accaggggtgg acgcttcgtc gcgcccaccc gtccacgaga ggaaccagca	360
atg acg ccg gcc atc tca cgc cgc gcc gtg ctc cag gga gcc ggc gcc Met Thr Pro Ala Ile Ser Arg Arg Ala Val Leu Gln Gly Ala Gly Ala	408
1 5 10 15	
gga gca ctc gcc ctg atc ttc ggc ggt gct gtg ccg cct gca gcc cgg Gly Ala Leu Ala Leu Ile Phe Gly Gly Ala Val Pro Pro Ala Ala Arg	456
20 25 30	
gca tcc gct ccg ggc tcg ctc cgt gcc gtc tac cac atg acg ccc ccc Ala Ser Ala Pro Gly Ser Leu Arg Ala Val Tyr His Met Thr Pro Pro	504
35 40 45	
agc ggc tgg ctc tgc gac ccc caa cgc ccg gtc acc acc cac ggc gcc Ser Gly Trp Leu Cys Asp Pro Gln Arg Pro Val Thr Thr His Gly Ala	552
50 55 60	
tac cag ctg tac tac ctg cac tcc gac cag aac aac ggc ccc ggc ggc Tyr Gln Leu Tyr Tyr Leu His Ser Asp Gln Asn Asn Gly Pro Gly Gly	600
65 70 75 80	
tgg gac cac gcg agc acg acc gac ggc gtc gcc ttc acg cac cac ggc Trp Asp His Ala Ser Thr Thr Asp Gly Val Ala Phe Thr His His Gly	648
85 90 95	
acc gtg atg ccg ctg cgc ccc gac ttc ccc gtg tgg tcc ggg tcg gcg Thr Val Met Pro Leu Arg Pro Asp Phe Pro Val Trp Ser Gly Ser Ala	696
100 105 110	
gtc gtc ggc acc gcg aac acg gca ggc ttc ggc gcc ggc gcg gtc gtc Val Val Gly Thr Ala Asn Thr Ala Gly Phe Gly Ala Gly Ala Val Val	744
115 120 125	
gcg ctc gcg acc cag ccg acc gac ggc gtc cgc aag tac cag gag cag Ala Leu Ala Thr Gln Pro Thr Asp Gly Val Arg Lys Tyr Gln Glu Gln	792
130 135 140	
tac ctc tac tgg tcg acc gac ggc ggc ttc acg ttc acc gcc ctg ccc Tyr Leu Tyr Trp Ser Thr Asp Gly Gly Phe Thr Phe Thr Ala Leu Pro	840
145 150 155 160	
gac ccc gtc atc gtc aac acc gac ggt cgc gcc gcc acc acg ccc gcc Asp Pro Val Ile Val Asn Thr Asp Gly Arg Ala Ala Thr Thr Pro Ala	888
165 170 175	
gag atc gag aac gcc gag tgg ttc cgc gac ccc aag atc cac tgg gac Glu Ile Glu Asn Ala Glu Trp Phe Arg Asp Pro Lys Ile His Trp Asp	936
180 185 190	
acc gcc cgc gga gaa tgg gtc tgc gtc atc gga cga ctg cgg tac gcc Thr Ala Arg Gly Glu Trp Val Cys Val Ile Gly Arg Leu Arg Tyr Ala	984
195 200 205	
gcg ttc tac acc tcg ccg aac ctg cgc gac tgg aca ctt cgc cgc aac Ala Phe Tyr Thr Ser Pro Asn Leu Arg Asp Trp Thr Leu Arg Arg Asn	1032
210 215 220	

sequence-976414[1].3(1).Oct02.ST25

ttc gac tac ccg aac cac gcc ctc ggc ggc atc gag tgc ccc gac ctg Phe Asp Tyr Pro Asn His 230 Ala Leu Gly Gly Ile Glu Cys Pro Asp Leu 240	1080
ttc gag atc acc gca gac gac ggg aca cgc cac tgg gtg ctc gcc gcc Phe Glu Ile Thr 245 Ala Asp Asp Gly Thr Arg His Trp Val Leu Ala Ala 255	1128
agc atg gac gcc tac ggc atc ggc ctc ccc atg acg tac gcc tac tgg Ser Met Asp Ala Tyr Gly Ile Gly Leu Pro Met Thr Tyr Ala Tyr Trp 270	1176
aca ggc acc tgg gac ggc gag cag ttc cac gcc gac gac ctc acc ccg Thr Gly Thr 275 Trp Asp Gly Glu Gln Phe His Ala Asp 285 Leu Thr Pro	1224
caa tgg ctc gac tgg ggc tgg gac tgg tac gcg gcc gtc acc tgg cca Gln Trp Leu Asp Trp Gly Trp 295 Asp Trp Tyr Ala Ala Val Thr Trp Pro 300	1272
tcg atc gac gcg ccc gag acc aag cgc ctc gcc atc gcg tgg atg aac Ser Ile Asp Ala Pro 310 Glu Thr Lys Arg Leu Ala Ile Ala Trp Met Asn 320	1320
aac tgg aag tac gcc gca cgc gac gtc ccc acc gac gca tcc gac ggc Asn Trp Lys Tyr 325 Ala Ala Arg Asp Val 330 Thr Asp Ala Ser 335 Asp Gly	1368
tac aac ggg cag aac tcg atc gtc cgc gag ctg cgg ctc gcc cga cag Tyr Asn Gly 340 Gln Asn Ser Ile Val Arg 345 Glu Leu Arg Leu Ala Arg Gln 350	1416
cct ggc ggc tgg tac acc ctc ctg agc acc ccc gtg gca gcg ctg acg Pro Gly 355 Gly Trp Tyr Thr Leu Leu Ser Thr Pro Val Ala Ala Leu Thr 365	1464
aac tac gtc acc gcc acc acc aca ctc ccc gac cgg acc gtc gac ggc Asn Tyr Val Thr Ala Thr 375 Thr Thr Leu Pro Asp Arg Thr Val Asp Gly 380	1512
agc gcc gtc ctg cca tgg aac gga cgc gca tac gag atc gag ctc gac Ser Ala Val Leu Pro 390 Trp Asn Gly Arg Ala Tyr 395 Glu Ile Glu Leu Asp 400	1560
atc gcc tgg gac acc gcg acg aac gtc ggc atc tcg gtg ggc cgc tcc Ile Ala Trp Asp Thr Ala Thr Asn Val Gly Ile Ser Val Gly Arg Ser 415	1608
ccc gac gga acc cgg cac acg aac atc ggc aag tac gga gca gac ctg Pro Asp Gly 420 Thr Arg His Thr Asn Ile Gly Lys Tyr Gly Ala Asp Leu 430	1656
tac gtc gac cga gga ccc tcc gac ctc gcc ggg tac tcg ctc gcc ccc Tyr Val Asp 435 Arg Gly Pro Ser Asp 440 Leu Ala Gly Tyr 445 Ser Leu Ala Pro	1704
tac tcg cga gcc gcc gcc ccc atc gac ccc ggc gcc cga tcc gtg cac Tyr Ser Arg Ala Ala Ala Pro 455 Ile Asp Pro Gly Ala Arg Ser Val His 460	1752
ctg cgc atc ctc gtc gac acc cag agc gtc gag gtc ttc gtc aac gcc Leu Arg Ile Leu Val Asp Thr Gln Ser Val Glu Val Phe Val Asn Ala	1800

465	sequence-976414[1].3(1).Oct02.ST25															480
ggc cac acc gtg ctc tcc cag cag gtc cac ttc gcc gag ggc gac acg	1848															
Gly His Thr Val Leu Ser Gln Gln Val His Phe Ala Glu Gly Asp Thr																
gga atc tcg ctc tac acc gac ggc ggc ccc gca cac ttc acc ggc atc	1896															
Gly Ile Ser Leu Tyr Thr Asp Gly Gly Pro Ala His Phe Thr Gly Ile																
gtc gtc cgc gag att ggc cag gcgatctagg cgatgcacac cacaccgctc	1947															
Val Val Arg Glu Ile Gly Gln																
accgaagccg cgccccggga gacgacggcc gacaatcgac acgtcctcgt cgtt	2001															

sequence-976414[1].3(1).Oct02.ST25

Asp Pro Val Ile Val Asn Thr Asp Gly Arg Ala Ala Thr Thr Pro Ala
165 170 175

Glu Ile Glu Asn Ala Glu Trp Phe Arg Asp Pro Lys Ile His Trp Asp
180 185 190

Thr Ala Arg Gly Glu Trp Val Cys Val Ile Gly Arg Leu Arg Tyr Ala
195 200 205

Ala Phe Tyr Thr Ser Pro Asn Leu Arg Asp Trp Thr Leu Arg Arg Asn
210 215 220

Phe Asp Tyr Pro Asn His Ala Leu Gly Gly Ile Glu Cys Pro Asp Leu
225 230 235 240

Phe Glu Ile Thr Ala Asp Asp Gly Thr Arg His Trp Val Leu Ala Ala
245 250 255

Ser Met Asp Ala Tyr Gly Ile Gly Leu Pro Met Thr Tyr Ala Tyr Trp
260 265 270

Thr Gly Thr Trp Asp Gly Glu Gln Phe His Ala Asp Asp Leu Thr Pro
275 280 285

Gln Trp Leu Asp Trp Gly Trp Asp Trp Tyr Ala Ala Val Thr Trp Pro
290 295 300

Ser Ile Asp Ala Pro Glu Thr Lys Arg Leu Ala Ile Ala Trp Met Asn
305 310 315 320

Asn Trp Lys Tyr Ala Ala Arg Asp Val Pro Thr Asp Ala Ser Asp Gly
325 330 335

Tyr Asn Gly Gln Asn Ser Ile Val Arg Glu Leu Arg Leu Ala Arg Gln
340 345 350

Pro Gly Gly Trp Tyr Thr Leu Leu Ser Thr Pro Val Ala Ala Leu Thr
355 360 365

Asn Tyr Val Thr Ala Thr Thr Thr Leu Pro Asp Arg Thr Val Asp Gly
370 375 380

Ser Ala Val Leu Pro Trp Asn Gly Arg Ala Tyr Glu Ile Glu Leu Asp
385 390 395 400

Ile Ala Trp Asp Thr Ala Thr Asn Val Gly Ile Ser Val Gly Arg Ser
405 410 415

sequence-976414[1].3(1).Oct02.ST25

Pro Asp Gly Thr Arg His Thr Asn Ile Gly Lys Tyr Gly Ala Asp Leu
420 425 430

Tyr Val Asp Arg Gly Pro Ser Asp Leu Ala Gly Tyr Ser Leu Ala Pro
435 440 445

Tyr Ser Arg Ala Ala Ala Pro Ile Asp Pro Gly Ala Arg Ser Val His
450 455 460

Leu Arg Ile Leu Val Asp Thr Gln Ser Val Glu Val Phe Val Asn Ala
465 470 475 480

Gly His Thr Val Leu Ser Gln Gln Val His Phe Ala Glu Gly Asp Thr
485 490 495

Gly Ile Ser Leu Tyr Thr Asp Gly Gly Pro Ala His Phe Thr Gly Ile
500 505 510

Val Val Arg Glu Ile Gly Gln
515